Kinect Depth Image Processing for Hand Motion Recognition using Backpropagation Neural Network

Syamsiar Kautsar[1], Purwadi A. Darwito[1], and Sryang T. Sarena[2]

Abstract-Kinect RGB-D Camera is a sensor for human gesture recognition. Kinect has built-in infrared array sensor which produces a depth image data. In this research, the depth image data is used to recognize human hand condition. The depth images of human right hand are processed by normalized central moment method. Furthermore, the moment value of hand depth image is used for backpropagation artificial neural network data training. A great training result is yielded by 70% accuracy. In addition, the output of neural netwok model is used to control a robot hand.

Index Terms – Kinect camera, depth image, neural network.

INTRODUCTION

In this research, a human hand motion recognition[1] system using Kinect RGB-D Camera[2] sensor is offered. This recognition system is used to control a robot hand motion [3][4]. The tracked object does not need to wear any pattern for recognition processing. The depth image of human hand is characterized by moment value. This value is extracted using normalized center moment method[5]. Hereafter, the moment value is processed using backpropagation artificial neural network (BP ANN)[6]. Hand grip condition is represented by the ANN output data. Finally, the result of hand motion recognition is used to put the BuTO robot hand in motion.

MATERIAL AND METHODS

I. The robot hand hardware

The shape and size of the robot hand is conditioned to grasp the objects with a diameter of 10 cm. Figure1 is the result of hardware realization of the robot hand. The robot hand actuators are using standard servo HS-255B model. The motions degree of servos are controlled through ATMega 16 which receives input data from computer via serial communication.

Figure 1. Robot Hand Hardware.

II. The characterization of depth image and BPANN modeling

The results of depth image cropping is based on 3D skeleton data, which then is processed using image clipping method to remove the background and change the grayscale image into bitmap image. The bitmap image is characterized using 1st and 2nd equations. The characterization results are used for ANN data input. In this study, the ANN model has 8 neurons in the 1st hidden layer, 4 neurons in the 2nd hidden layer, and tangen sigmoid as activation function. Figure 2 is an example of the depth image processing for hand motion recognition.

$$M_{ij} = \sum_{x=1}^{M}\sum_{y=1}^{N}(x-\bar{x})^i(y-\bar{y})^j.I(x,y). \quad (1)$$

$$\eta_{ij} = \frac{\mu_{pq}}{\mu^{\gamma}_{00}}, \gamma = \frac{i+j+2}{2} \quad (2)$$

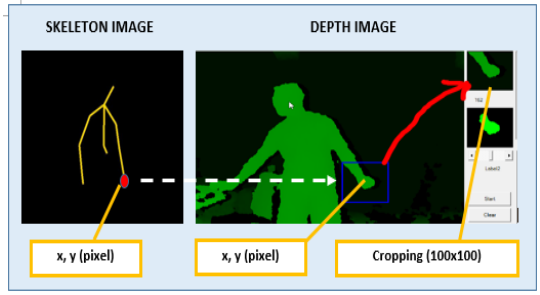

Figure 2. Depth Image Processing.

Each hand condition has different N11, N02, N20, N03, N30, N12, N21 value. The moment values which are processed by ANN represent the human hand condition. The system reads it as opened hand condition if ANN output value < 0.55, and closed hand condition if ANN output value >=0.55.

[1]Syamsiar Kautsar and Purwadi A. Darwito are with Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia. E-mail: kautsar.sam@gmail.com; padarwito@yahoo.com.
[2]Sryang T. Sarena is with Politeknik Elektronika Negeri, Surabaya, Indonesia. E-mail: sryang.e@gmail.com.

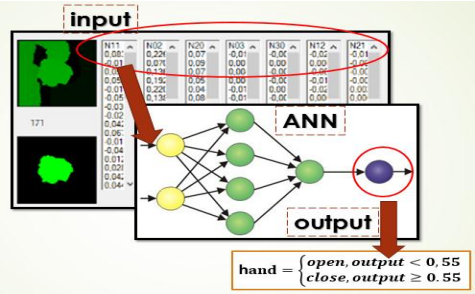

Figure 3. ANN for human hand recognition.

RESULT AND DISCUSSION

For the hand recognition testing, the output of ANN is stored in PC every 50ms for 6.25 seconds. As the result of closed hand condition test, 96 values of 125 data are accurately above the limit value (5.5). So the percentage of successful hand recognition system is 76.8%. Figure 4 is the ANN data output of the first test.

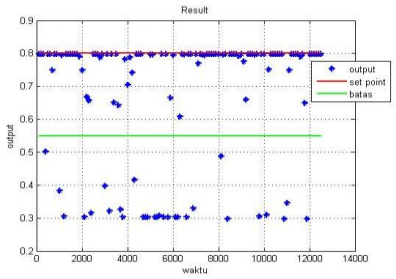

Figure 4. ANN output for first test.

The second test is performed on the condition of the moving arm. In opened hand condition, 91 values of the 125 result data are below the limit value (5.5). Hence, the percentage of the hand recognition system accuracy is 72.8%. The ANN data output for the second test is shown by Figure 5.

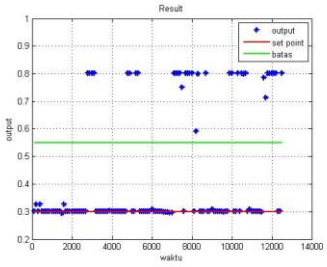

Figure 5. ANN output for SECOND test.

CONCLUSION

In this study, a BPANN method is used for human hand recognition based on depth image of Kinect camera. The accuracy rate of the hand recognition is more than 70%. For practical implementation, the ANN output can be used to control the robot hand for grasping an object.

ACKNOWLEDGEMENTS

Our gratitude goes to DIKTI, which has awarded fresh graduate scholarships to the author for the post-graduate program in Physics Engineering ITS.

REFERENCES

[1] Ren, Zhou, et all., "Robust Part-Based Hand Gesture Recognition Using Kinect Sensor", *IEEE Transactions on Multimedia*, Vol. 15, No. 5, pg. 1110-1120, 2013.

[2] Tashev, Ivan, "Kinect Development Kit: A Toolkit for Gesture-and Speech-Based Human-Machine Interaction", *IEEE Signal Processing Magazine*, 2013, pg. 129-131.

[3] Megalingam, R.K., et all., "Kinect Based Gesture Controlled Robotic Arm: A research work at HuT Labs", *IEEE International Conference in MOOC, Innovation and Technology in Education* (MITE), pg. 294-299.

[4] Kofman, Jonathan, Wu, Xianghai, Luu, T.J., and Verma, Siddharth, "Teleoperation of a Robot Manipulator using a Vision-Based Human-Robot Interface", IEEE Transactions on Industrial Electronics, Vol. 52, No. 5, pg. 1206-1219,2005.

[5] Kadir, Abdul, *Dasar Pengolahan Citra dengan Delphi*, Yogyakarta: Penerbit Andi, 2013.

[6] Purnomo, Hery Mauridhi and Kurniawan, Agus, *Supervised Neural Networks dan Aplikasinya*, Yogyakarta: Graha Ilmu, 2006.

Syamsiar Kautsar, obtained his Bachelor of Applied Science (S.ST.), Automation Engineering Study Program in Shipbuilding State Polytechnic of Surabaya (SHIPS), and had graduated in 2013. Currently, he is a laboratory assistant in the Automation and Robotic Laboratory - SHIPS and pursuing his Master degree in Physics Engineering Department of Sepuluh Nopember Institute of Technology.